The elimination matrix

Magnus, J.R.; Neudecker, H.

Published in:
SIAM Journal on Algebraic and Discrete Methods

Publication date:
1980

SIAM J. ALG. DISC. METH.
Vol. 1, No. 4, December 1980

0196-5212/80/0104-0008 $01.00/0

THE ELIMINATION MATRIX: SOME LEMMAS AND APPLICATIONS*

JAN R. MAGNUS† AND H. NEUDECKER‡

Abstract. Two transformation matrices are introduced, L and D, which contain zero and unit elements only. If A is an arbitrary (n,n) matrix, L eliminates from $\mathrm{vec}A$ the supradiagonal elements of A, while D performs the inverse transformation for *symmetric* A. Many properties of L and D are derived, in particular in relation to Kronecker products. The usefulness of the two matrices is demonstrated in three areas of mathematical statistics and matrix algebra: maximum likelihood estimation of the multivariate normal distribution, the evaluation of Jacobians of transformations with symmetric or lower triangular matrix arguments, and the solution of matrix equations.

1. Introduction. If a matrix A has a known structure (symmetric, skew symmetric, diagonal, triangular), some elements of A are redundant in the sense that they can be deduced from this structure. Thus, if A is a symmetric or lower triangular matrix of order n, its $\frac{1}{2}n(n-1)$ supradiagonal elements are redundant. If we eliminate these elements from $\mathrm{vec}A$ (the column vector stacking the columns of A), this defines a new vector of order $\frac{1}{2}n(n+1)$ which we denote as $v(A)$. The matrix which, for arbitrary A, transforms $\mathrm{vec}A$ into $v(A)$ is the elimination matrix L, first mentioned by Tracy and Singh (1972) and later by Vetter (1975) and Balestra (1976).

Of equal interest is the inverse transformation from $v(A)$ to $\mathrm{vec}A$. For lower triangular A, we shall see that $L'v(A)=\mathrm{vec}A$. We further introduce the duplication matrix D such that, for symmetric A, $Dv(A)=\mathrm{vec}A$. The matrix D (or a matrix comparable to D) was previously defined by Tracy and Singh (1972), Browne (1974), Vetter (1975), Balestra (1976), and Nel (1978). D^+, the Moore-Penrose inverse of D, possesses the property, used by Browne (1974) and Nel (1978), $D^+\mathrm{vec}A=v(A)$, for *symmetric* A.

The purpose of this paper is to study the matrices L and D. Both matrices consist of zero and unit elements only. §2 gives the necessary definitions and basic tools. The next two sections contain the theoretical heart of the paper and establish a number of results on L and D. §§5–7 are devoted to applications: maximum likelihood estimation of the multivariate normal distribution, the evaluation of Jacobians of transformations with symmetric or lower triangular matrix arguments, and, finally, the solution of matrix equations. An appendix presents the proofs of the lemmas in §4.

Not all results are new. Thus, Tracy and Singh (1972) established that $|L(A\otimes A)D|=|A|^{n+1}$. They obtained two other determinants as well (their examples 5.3 and 5.4), but these are both in error. Browne (1974) proved the important fact that $(D'(A\otimes A)D)^{-1}=D^+(A^{-1}\otimes A^{-1})D^{+\prime}$ for nonsingular A, while Nel (1978) evaluated the determinant of $D^+(A\otimes B)D$, when $AB=BA$, A and B symmetric. Concurrently with the present paper, Henderson and Searle (1979) wrote an article on the same topic. Inevitably there is some overlap between the two papers.

2. Notation and preliminary results. All matrices are real; capital letters represent matrices; lowercase letters denote vectors or scalars. An (m,n) matrix is one having m

*Received by the editors September 26, 1978 and in final form February 7, 1980.

†Department of Economics, University of British Columbia, 2075 Wesbrook Mall, Vancouver, Canada V6T 1W5.

‡Institute of Actuarial Science and Econometrics, University of Amsterdam, 23 Jodenbreestraat, 1011 NH Amsterdam, Netherlands.

rows and n columns; A' denotes the transpose of A, $\mathrm{tr}A$ its trace, and $|A|$ its determinant. If A is a square matrix, $\bar{A}$ denotes the lower triangular matrix derived from A by setting all supradiagonal elements in A equal to zero; $\mathrm{dg}(A)$ is the diagonal matrix derived from A by setting all supra- *and* infradiagonal elements in A equal to zero. If x is an n-vector, and $f(x)=(f_1(x)\cdots f_m(x))'$ a differentiable vector function of x, then the matrix $\partial f/\partial x$ has order (n,m) with typical element $(\partial f_j/\partial x_i)$.

The unit vector e_i, $i=1,\cdots,n$, is the ith column of the identity matrix I_n, i.e., it is an n-vector with one in its ith position and zeroes elsewhere. The (n,n) matrix E_{ij} has one in its ijth position and zeroes elsewhere, i.e., $E_{ij}=e_ie_j'$. We partition the identity matrix of order $\frac{1}{2}n(n+1)$ as follows:

$$I_{(1/2)n(n+1)}=(u_{11}u_{21}\cdots u_{n1}u_{22}\cdots u_{n2}u_{33}\cdots u_{nn}).$$

Formally, u_{ij} is a unit vector of order $\frac{1}{2}n(n+1)$ with unity in its $[(j-1)n+i-\frac{1}{2}j(j-1)]$-th position and zeroes elsewhere $(1\leqq j\leqq i\leqq n)$.

If A is an (m,n) matrix and $A_{\cdot j}$ its jth column, then $\mathrm{vec}A$ is the mn-column vector

$$\mathrm{vec}A=\begin{bmatrix}A_{\cdot 1}\\ \vdots\\ A_{\cdot n}\end{bmatrix}.$$

If A is square of order n, $v(A)$ denotes the $\frac{1}{2}n(n+1)$ vector that is obtained from $\mathrm{vec}A$ by eliminating all supradiagonal elements of A. For example, if $n=3$,

$$\mathrm{vec}A=(a_{11}a_{21}a_{31}a_{12}a_{22}a_{32}a_{13}a_{23}a_{33})',$$

and

$$v(A)=(a_{11}a_{21}a_{31}a_{22}a_{32}a_{33})'.$$

Finally, the Kronecker product of an (m,n) matrix $A=(a_{ij})$ and an (s,t) matrix B is the (ms,nt) matrix

$$A\otimes B=(a_{ij}B).$$

This settles the notation. Let us now state some preliminary results that will be used throughout. If $A=(a_{ij})$ is an (n,n) matrix, then A, $\bar{A}$, and $\mathrm{dg}(A)$ can be expressed as

$$A=\sum_{ij}a_{ij}E_{ij};\qquad \bar{A}=\sum_{i\geqq j}a_{ij}E_{ij};\qquad \mathrm{dg}(A)=\sum_{i=1}^{n}a_{ii}E_{ii}. \tag{2.1}$$

A standard result on vecs is

$$\mathrm{vec}ABC=(C'\otimes A)\mathrm{vec}B, \tag{2.2}$$

if the matrix product ABC exists. For vectors x and y of any order we then have

$$x\otimes y=\mathrm{vec}\,yx' \qquad\text{and}\qquad x\otimes y'=xy'=y'\otimes x. \tag{2.3}$$

The basic connection between the vec-function and the trace is

$$(\mathrm{vec}A)'\mathrm{vec}B=\mathrm{tr}A'B, \tag{2.4}$$

where A and B are (m,n) matrices. From (2.2) and (2.4) follows

$$(\mathrm{vec}A)'(B\otimes C)\mathrm{vec}D=\mathrm{tr}A'CDB', \tag{2.5}$$

if the expression on the right-hand side exists.

We shall frequently use the commutation matrix K defined implicitly as:

DEFINITION 2.1a (implicit definition of K). The (n^2, n^2) *commutation matrix* K performs for every (n, n) matrix A the transformation $K\operatorname{vec}A = \operatorname{vec}A'$.

In fact, K is a special case of the (mn, mn) matrix K_{mn} which maps $\operatorname{vec}A$ into $\operatorname{vec}A'$ for an arbitrary (m, n) matrix A. The matrix K_{mn} was introduced by Tracy and Dwyer (1969). Many of its properties are derived in Magnus and Neudecker (1979), who also established the following explicit expression for K.

DEFINITION 2.1b (explicit definition of K).

$$K = \sum_{i=1}^{n} \sum_{j=1}^{n} (E_{ij} \otimes E'_{ij}).$$

Closely related to the commutation matrix is the matrix N.

DEFINITION 2.2a (implicit definition of N). The (n^2, n^2) matrix N performs for every (n, n) matrix A the transformation $N\operatorname{vec}A = \operatorname{vec}\frac{1}{2}(A + A')$.

Its explicit expression is immediately derived.

DEFINITION 2.2b (explicit definition of N).

$$N = \tfrac{1}{2}(I + K).$$

Note that the implicit definitions of K and N are proper definitions in the sense that they uniquely determine K and N. The following lemma gives some properties of K and N.

LEMMA 2.1.

(i) $K = K' = K^{-1}$;

(ii) $K(A \otimes B) = (B \otimes A)K$, *for any* (n, n) *matrices* A *and* B;

(iii) $N = N' = N^2$;

(iv) $NK = N = KN$.

For any (n, n) *matrix* A *we have*

(v) $N(A \otimes A) = (A \otimes A)N = N(A \otimes A)N$;

(vi) $N(I \otimes A + A \otimes I) = (I \otimes A + A \otimes I)N = N(I \otimes A + A \otimes I)N$
$= 2N(I \otimes A)N = 2N(A \otimes I)N$.

Proof. The properties of K follow from Magnus and Neudecker (1979). The properties of N follow from those of K since $N = \frac{1}{2}(I + K)$. □

Let us now give four results on the unit vector u_{ij} of order $\frac{1}{2}n(n+1)$ and the $v(\cdot)$ operator.

$$\sum_{i \geq j} u_{ij} u'_{ij} = I_{(1/2)n(n+1)}. \tag{2.6}$$

If A is an (n, n) matrix, then

$$v(A) = v(\bar{A}) = \sum_{i \geq j} a_{ij} u_{ij} \quad \text{and} \quad v(\operatorname{dg}(A)) = \sum_{i} a_{ii} u_{ii}; \tag{2.7}$$

$$u_{ij} = v(E_{ij}) \quad \text{and} \quad a_{ij} = u'_{ij} v(A), \qquad i \geq j; \tag{2.8}$$

$$v(A) = v(\operatorname{dg}(A)), \text{ if } A \text{ is upper triangular.} \tag{2.9}$$

Finally, we make use of the following standard facts in matrix differentiation. For every matrix X and Y of appropriate orders,

$$d(XY) = (dX)Y + X(dY), \tag{2.10}$$

$$d\operatorname{tr}XY = \operatorname{tr}(dX)Y + \operatorname{tr}X dY. \tag{2.11}$$

For every nonsingular X,

$$dlog|X| = \mathrm{tr}X^{-1}dX, \tag{2.12}$$

$$dX^{-1} = -X^{-1}(dX)X^{-1}. \tag{2.13}$$

3. Basic properties of L and D. Let us now introduce the elimination matrix L. As in the previous section, where we defined K and N, the elimination matrix will be defined implicitly *and* explicitly.

DEFINITION 3.1a (implicit definition of L). The $(\frac{1}{2}n(n+1), n^2)$ *elimination matrix* L performs for every (n, n) matrix A the transformation $L\mathrm{vec}A = v(A)$.

L, thus defined, eliminates from $\mathrm{vec}A$ the supradiagonal elements of A. We shall show that L is uniquely determined by (3.1a). Let A be an arbitrary (n, n) matrix, and suppose that $\tilde{L}$ and L both transform $\mathrm{vec}A$ into $v(A)$. Then $(\tilde{L}-L)\mathrm{vec}A = 0$ for every A. Hence, $\tilde{L} = L$. We can derive an explicit expression for L as follows. Recall that e_i, $i = 1 \cdots n$, is the ith unit vector of order n. Then, using (2.7), (2.4), and (2.3), we find

$$\begin{aligned} v(A) &= \sum_{i \geqq j} a_{ij}u_{ij} = \sum_{i \geqq j} u_{ij}(e_i'Ae_j) = \sum_{i \geqq j} u_{ij}\mathrm{tr}(e_je_i'A) \\ &= \sum_{i \geqq j} u_{ij}\mathrm{tr}(E_{ij}'A) = \sum_{i \geqq j} u_{ij}(\mathrm{vec}\,E_{ij})'\mathrm{vec}A = \sum_{i \geqq j} (u_{ij} \otimes e_j' \otimes e_i')\mathrm{vec}A. \end{aligned}$$

This leads to the following explicit definition.

DEFINITION 3.1b (explicit definition of L).

$$L = \sum_{i \geqq j} u_{ij}(\mathrm{vec}\,E_{ij})' = \sum_{i \geqq j} (u_{ij} \otimes e_j' \otimes e_i').$$

An example, for $n = 3$, is

$$L = \left[\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & & & & & & \\ 0 & 1 & 0 & & 0 & & & 0 & \\ 0 & 0 & 1 & & & & & & \\ \hline & 0 & & 0 & 1 & 0 & & 0 & \\ & & & 0 & 0 & 1 & & & \\ \hline & 0 & & & 0 & & 0 & 0 & 1 \end{array}\right].$$

Most authors on (0, 1) matrices are interested only in transformations with *symmetric* matrices, and work with LN rather than L. See, e.g., Browne (1974) and Nel (1978). The justification for this lies in the following lemma.

LEMMA 3.1. *For any (n, n) matrix A we have*

(i) $$LN\mathrm{vec}A = \frac{1}{2}v(A + A').$$

In particular, when A is symmetric,

(ii) $$LN\mathrm{vec}A = v(A).$$

Proof. Immediate from the implicit definitions of N and L. □

Thus, if A is symmetric, L and LN play the same role. In this paper we have chosen a more general approach, based on $L\mathrm{vec}A = v(A)$ for *arbitrary* A, largely because this allows us to study transformations with triangular matrices as well. The following lemma characterizes L as a (0, 1) matrix with $\frac{1}{2}n(n+1)$ 1's, one in each row and not more than one in each column.

LEMMA 3.2.

(i) *L has full row-rank* $\frac{1}{2}n(n+1)$;

(ii) $LL'=I_{(1/2)n(n+1)}$;

(iii) $L^+=L'$, *where* L^+ *is the Moore-Penrose generalized inverse of* L.

Proof. We shall show that $LL'=I$. The other two results then follow directly.

$$LL'=\sum_{i\geq j}(u_{ij}\otimes e_j'\otimes e_i')\sum_{h\geq k}(u_{hk}'\otimes e_k\otimes e_h)=\sum_{i\geq j}\sum_{h\geq k}(u_{ij}u_{hk}'\otimes e_j'e_k\otimes e_i'e_h)$$

$$=\sum_{i\geq j}u_{ij}u_{ij}'=I_{(1/2)n(n+1)}, \qquad \text{by (2.6).} \quad \square$$

Let us now determine three matrices that are useful for certain linear transformations.

LEMMA 3.3. *The matrices* $L'L$, LKL', *and* $L'LKL'L$ *are diagonal and idempotent of rank* $\frac{1}{2}n(n+1)$, n, *and* n *respectively. Let* A *be an arbitrary* (n,n) *matrix. Then,*

(i) $L'L\operatorname{vec}A=\operatorname{vec}\bar{A}$;

(ii) $L'L=\sum_{i\geq j}(E_{jj}\otimes E_{ii})$;

(iii) $LKL'v(A)=v(\operatorname{dg}(A))$;

(iv) $LKL'=\sum_{i=1}^{n}u_{ii}u_{ii}'$;

(v) $L'LKL'L\operatorname{vec}A=\operatorname{vec}(\operatorname{dg}(A))$;

(vi) $L'LKL'L=\sum_{i=1}^{n}(E_{ii}\otimes E_{ii})$.

Proof. By the explicit definition of L we have

$$L'L=\sum_{i\geq j}(u_{ij}'\otimes e_j\otimes e_i)\sum_{h\geq k}(u_{hk}\otimes e_k'\otimes e_h')=\sum_{i\geq j}\sum_{h\geq k}(u_{ij}'u_{hk}\otimes e_je_k'\otimes e_ie_h')$$

$$=\sum_{i\geq j}(e_je_j'\otimes e_ie_i')=\sum_{i\geq j}(E_{jj}\otimes E_{ii}),$$

so that, using (2.2) and (2.1),

$$L'L\operatorname{vec}A=\sum_{i\geq j}(E_{jj}\otimes E_{ii})\operatorname{vec}A=\sum_{i\geq j}\operatorname{vec}(E_{ii}AE_{jj})$$

$$=\operatorname{vec}\sum_{i\geq j}(e_ie_i'Ae_je_j')=\operatorname{vec}\sum_{i\geq j}a_{ij}E_{ij}=\operatorname{vec}\bar{A}.$$

Further, for arbitrary $v(A)$,

$$LKL'v(A)=LK\operatorname{vec}\bar{A}=L\operatorname{vec}\bar{A}'=v(\bar{A}')=v(\operatorname{dg}(A))$$

$$=\sum_i a_{ii}u_{ii}=\sum_i u_{ii}u_{ii}'v(A),$$

by (i), the implicit definitions of K and L, (2.9), (2.7) and (2.8). This proves (iii) and (iv). Similarly, for arbitrary $\operatorname{vec}A$,

$$L'LKL'L\operatorname{vec}A=L'v(\operatorname{dg}(A))=\operatorname{vec}(\operatorname{dg}(A))=\operatorname{vec}\sum_i(a_{ii}E_{ii})$$

$$=\operatorname{vec}\sum_i(e_ie_i'Ae_ie_i')=\sum_i\operatorname{vec}(E_{ii}AE_{ii})=\sum_i(E_{ii}\otimes E_{ii})\operatorname{vec}A,$$

by (iii), (i), (2.1) and (2.2). It is easy to see that the three matrices are diagonal with only zeroes and ones on the diagonal. Hence they are idempotent. The rank of each of the three matrices equals the number of ones on the diagonal, i.e., $\frac{1}{2}n(n+1)$, n, and n respectively. □

Note that Lemma 3.3(i) implies that $L'L\text{vec}A=\text{vec}A$ if and only if A is lower triangular. The matrix LKL', as shown in the previous lemma, is diagonal with n ones and $\frac{1}{2}n(n-1)$ zeroes. Hence, $I+LKL'$ is a nonsingular diagonal matrix with n times 2 and $\frac{1}{2}n(n-1)$ times 1 on the diagonal. Because $LNL'=\frac{1}{2}L(I+K)L'=\frac{1}{2}(LL'+LKL')=\frac{1}{2}(I+LKL')$, it is diagonal too with n times 1 and $\frac{1}{2}n(n-1)$ times $\frac{1}{2}$ on the diagonal. The following properties of LNL' are of interest.

LEMMA 3.4. *The matrix LNL' is diagonal with determinant*

(i) $|LNL'|=2^{-(1/2)n(n-1)}$.

Its inverse is

(ii) $(LNL')^{-1}=2I-LKL'$.

Proof. Since LNL' is a diagonal matrix, its determinant is the product of its diagonal elements, i.e., $|LNL'|=2^{-(1/2)n(n-1)}$. Property (ii) is easily established using $LNL'=\frac{1}{2}(I+LKL')$ and the idempotency of LKL'. □

As we have seen, L uniquely transforms $\text{vec}A$ into $v(A)$. The inverse transformation generally does not exist. We can, however, easily transform $v(A)$ into (the vecs of) a lower triangular matrix or a diagonal matrix, since

$$L'v(A)=\text{vec}\bar{A} \qquad \text{(Definition 3.1a and Lemma 3.3 (i))},$$

and

$$L'LKL'v(A)=\text{vec}\,\text{dg}(A) \qquad \text{(Definition 3.1a and Lemma 3.3(v))}.$$

Combining these two transformations one verifies that

$$(L'+KL'-L'LKL')v(A)=\text{vec}\big(\bar{A}+\bar{A}'-\text{dg}(A)\big).$$

We have thus found a matrix which transforms $v(A)$ into (the vec of) a *symmetric* matrix. Let us define this matrix implicitly.

DEFINITION 3.2a (implicit definition of D). The $(n^2,\frac{1}{2}n(n+1))$ *duplication matrix* D performs for every (n,n) matrix A the transformation $Dv(A)=\text{vec}(\bar{A}+\bar{A}'-\text{dg}(A))$.

It is easy to see that D is unique. Hence, $D=L'+KL'-L'LKL'=2NL'-L'LKL'$. Note that in particular, if A is symmetric, $DL\text{vec}A=Dv(A)=\text{vec}A$. This is an important property that we will frequently use. The converse is also true; i.e., any A satisfying $DL\text{vec}A=\text{vec}A$ is symmetric.

LEMMA 3.5.

(i) $LD=I_{(1/2)n(n+1)}$;

(ii) $DLN=N$;

(iii) $D=2NL'-L'LKL'=NL'(LNL')^{-1}$.

Proof. Let $A=A'$; then $LDv(A)=L\text{vec}A=v(A)$. Hence, $LD=I$, since the symmetry of A does not restrict $v(A)$. Further, for arbitrary A,

$$DLN\text{vec}A=DL\text{vec}\tfrac{1}{2}(A+A')=Dv\big(\tfrac{1}{2}(A+A')\big)=\text{vec}\tfrac{1}{2}(A+A')=N\text{vec}A,$$

which proves (ii). It also implies that $DLNL'=NL'$, and because of the nonsingularity of LNL', $D=NL'(LNL')^{-1}$. □

Note that $DLN=N$ is a defining property of D. In fact, it is just a reformulation of Definition 3.2a. The matrix D can be explicitly expressed in terms of unit vectors of order $\frac{1}{2}n(n+1)$ and n, i.e., in terms of u_{ij}, e_i, and e_j. From the explicit definition of K and L, and the expression for $L'L$ (Lemma 3.3 (ii)), one verifies that

$$LK = \sum_{i \geq j} u_{ij}(\text{vec}\, E_{ji})',$$

and

$$LKL'L = \sum_{i} u_{ii}(\text{vec}\, E_{ii})',$$

so that

$$\begin{aligned} D' &= L + LK - LKL'L \\ &= \sum_{i \geq j} u_{ij}(\text{vec}\, E_{ij})' + \sum_{i \geq j} u_{ij}(\text{vec}\, E_{ji})' - \sum_{i} u_{ii}(\text{vec}\, E_{ii})'. \end{aligned}$$

Hence, we may define D as follows.

DEFINITION 3.2b (explicit definition of D). Let T_{ij} be an (n,n) matrix with 1 in its ijth and jith position, and zeroes elsewhere. Then

$$D' = \sum_{i \geq j} u_{ij}(\text{vec}\, T_{ij})'.$$

Note that $T_{ij} = E_{ij} + E_{ji}$ for $i \neq j$, and that $T_{ii} = E_{ii}$. An example, for $n=3$, is

$$D = \left[\begin{array}{ccc|ccc|c} 1 & 0 & 0 & & & & \\ 0 & 1 & 0 & & 0 & & 0 \\ 0 & 0 & 1 & & & & \\ \hline 0 & 1 & 0 & 0 & 0 & & \\ 0 & 0 & 0 & 1 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 1 & & \\ \hline 0 & 0 & 1 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 1 & & 0 \\ 0 & 0 & 0 & 0 & 0 & & 1 \end{array}\right].$$

Further properties of D are contained in the following two lemmas.

LEMMA 3.6.

(i) *D has full column-rank* $\frac{1}{2}n(n+1)$;

(ii) $KD = D = ND$;

(iii) $D'D = (LNL')^{-1}$;

(iv) $D^{+} = LN$.

Proof. Straightforward from the expression $D = NL'(LNL')^{-1}$ and the properties $DLN = N = N^2 = KN$, and $LD = I$. □

LEMMA 3.7. *Let A be an arbitrary (n,n) matrix. Then,*

(i) $D'\text{vec}A = v(A + A' - \text{dg}(A))$;

(ii) $DD'\text{vec}A = \text{vec}(A + A' - \text{dg}(A))$;

(iii) $DD' = 2N - L'LKL'L$.

Proof. From Lemmas 3.3(iii) and (v), and 3.5(iii), and Definitions 2.1a, 2.2b, 3.1a, and 3.2a, we have

$$D'\text{vec}A = (L + LK - LKL'L)\text{vec}A = v(A) + v(A') - v(\text{dg}(A)),$$

and

$$DD'\text{vec}A = Dv(A+A'-\text{dg}(A)) = \text{vec}(A+A'-\text{dg}(A))$$
$$= (I+K)\text{vec}A - L'LKL'L\text{vec}A = (2N - L'LKL'L)\text{vec}A,$$

for an arbitrary (n,n) matrix A. Hence, $DD' = 2N - L'LKL'L$. □

The matrices L and D, like the commutation matrix K, are useful in matrix differentiation. From Definition 2.1a, it follows that $\partial\,\text{vec}X/\partial\,\text{vec}X' = K$ where X is an (n,n) matrix. The corresponding results for L and D are contained in the following lemma.

LEMMA 3.8. *Let X be an (n,n) matrix. Then*

(i) $\partial\text{vec}X/\partial v(X) = \begin{cases} L, & \textit{for lower triangular } X; \\ D', & \textit{for symmetric } X; \end{cases}$

(ii) $\partial v(X)/\partial\text{vec}X = L'$.

Proof. Immediate from the relations $\text{vec}X = L'v(X)$ (lower triangular X), $\text{vec}X = Dv(X)$ (symmetric X), and $v(X) = L\text{vec}X$. □

Comment. More general results are easily obtained from Lemma 3.8, using the chain rule. In particular, let $Y = F(x)$ be an (n,n) matrix whose elements are differentiable functions of a vector x. Then

(i) $\partial\,\text{vec}Y/\partial x = \begin{cases} (\partial v(Y)/\partial x)L & \text{if } Y \text{ is lower triangular for all } x; \\ (\partial v(Y)/\partial x)D' & \text{if } Y \text{ is symmetric for all } x; \end{cases}$

(ii) $\partial v(Y)/\partial x = (\partial\,\text{vec}Y/\partial x)L'$ for all Y.

4. Applications to Kronecker products. From §2 we know that the commutation matrix K possesses two major properties: a *transformation* property, $K\text{vec}A = \text{vec}A'$ (its definition), and a *Kronecker* property, $K(A\otimes B)K = B\otimes A$. The elimination matrix L and the duplication matrix D have likewise been defined by their transformation properties, viz. $L\text{vec}A = v(A)$ and, for symmetric A, $Dv(A) = \text{vec}A$. Let us now investigate their Kronecker properties. The applications in §§5–7 are based almost entirely on the lemmas in the present section. Proofs are postponed to the Appendix.

We shall first show that, if A and B have a certain structure (diagonal, triangular), Kronecker forms of the type $L(A\otimes B)L'$ and $L(A\otimes B)D$ often possess the same structure.

LEMMA 4.1. *Let Λ and M be diagonal n-matrices with diagonal elements λ_i and μ_i $(i=1\cdots n)$. Let further P and Q be lower triangular n-matrices with diagonal elements p_{ii} and q_{ii}, $i=1\cdots n$. Then,*

(i) *$L(\Lambda\otimes M)L' = L(\Lambda\otimes M)D$ is diagonal with elements $\mu_i\lambda_j (i\geqq j)$ and determinant $\Pi_i\mu_i^i\lambda_i^{n-i+1}$;*

(ii) *$L(P\otimes Q)L'$ is lower triangular with diagonal elements $q_{ii}p_{jj}(i\geqq j)$ and determinant $\Pi_i q_{ii}^i p_{ii}^{n-i+1}$;*

(iii) *$L(P\otimes Q)D$ is lower triangular and $L(P'\otimes Q')D$ is upper triangular. Both matrices have diagonal elements $q_{ii}p_{jj}(i\geqq j)$ and determinant $\Pi_i q_{ii}^i p_{ii}^{n-i+1}$.*

Next we establish some properties of $L(P'\otimes Q)L'$, with lower triangular P and Q. Notice that (i) is the Kronecker counterpart of the property $L'L\text{vec}P = \text{vec}P$ for lower triangular P.

LEMMA 4.2. *For lower triangular n-matrices $P = (p_{ij})$ and $Q = (q_{ij})$,*

(i) $L'L(P'\otimes Q)L' = (P'\otimes Q)L'$;

(ii)

$$[L(P'\otimes Q)L']^s = L[(P')^s\otimes Q^s]L', \quad \begin{cases} s=0,1,2,\cdots, \\ s=\cdots,-2,-1, & \text{if } P^{-1} \text{ and } Q^{-1} \text{ exist}, \\ s=\frac{1}{2}, & \text{if lower triangular } P^{1/2} \text{ and } Q^{1/2} \text{ exist}; \end{cases}$$

(iii) $L(P'\otimes Q)L' = D'(P'\otimes Q)L'$ *has eigenvalues* $q_{ii}p_{jj}$ $(i \geqq j)$ *and determinant* $\Pi_i q_{ii}^i p_{ii}^{n-i+1}$.

In Lemma 4.1 we have proved that, for lower triangular P and Q, the matrices $L(P\otimes Q)L'$, $L(P\otimes Q)D$, $L(P'\otimes Q')L'$, and $L(P'\otimes Q')D$ are triangular as well, with diagonal elements $q_{ii}p_{jj}$, $i\geqq j$. Although the matrices $L(P'\otimes Q)L'$, $L(P\otimes Q')L'$, and $L(P\otimes Q')D$ are not triangular, they also possess eigenvalues $q_{ii}p_{jj}$, $i\geqq j$; see Lemma 4.2. The matrix $L(P'\otimes Q)D$ is more complicated and seems not to have such nice properties. In particular, its eigenvalues are in general different from $q_{ii}p_{jj}$, $i\geqq j$.

The results of Lemmas 4.1 and 4.2 enable us to find the following determinants which are of importance in the evaluation of Jacobians of transformations with lower triangular matrix arguments (see §6).

LEMMA 4.3. *For lower triangular n-matrices P, Q, R, and S with diagonal elements p_{ii}, q_{ii}, r_{ii}, and s_{ii}, $i=1\cdots n$, we have*

(i) $|L(PQ'\otimes R'S)L'| = \Pi_i(r_{ii}s_{ii})^i(p_{ii}q_{ii})^{n-i+1}$;

(ii) $|L(P'\otimes Q+R'\otimes S)L'| = \Pi_{i\geqq j}(q_{ii}p_{jj}+s_{ii}r_{jj})$;

(iii) $|L(P\otimes Q'R)D| = \Pi_i(q_{ii}r_{ii})^i p_{ii}^{n-i+1}$;

(iv) $|L(PQ'\otimes R')D| = \Pi_i r_{ii}^i(p_{ii}q_{ii})^{n-i+1}$.

If P, Q, R, S are nonsingular,

(v) $[L(PQ'\otimes R'S)L']^{-1} = L(Q'^{-1}\otimes S^{-1})L'L(P^{-1}\otimes R'^{-1})L'$.

Finally,

(vi) $$\left|L\sum_{h=1}^{H}[(P')^{H-h}\otimes P^{h-1}]L'\right| = H^n|P|^{H-1}\Pi_{i>j}\mu_{ij}, \qquad H=2,3,\cdots,$$

where

$$\mu_{ij} = \begin{cases} \dfrac{(p_{ii}^H - p_{jj}^H)}{(p_{ii}-p_{jj})}, & \text{if } p_{ii}\neq p_{jj}, \\ Hp_{ii}^{H-1}, & \text{if } p_{ii}=p_{jj}. \end{cases}$$

A variety of corollaries flow from Lemma 4.3 by putting one or more of the matrices P, Q, R, and S equal to I. Also, the four matrices $L(PQ'\otimes Q'P)L'$, $L(PQ'\otimes P'Q)L'$, $L(PQ\otimes Q'P)D$, and $L(PQ'\otimes P'Q')D$ have the same determinant, namely $|P|^{n+1}|Q|^{n+1}$.

In Lemmas 4.1–4.3 we have studied triangular matrices only. The crucial properties for lower triangular matrices are $L'L\operatorname{vec}P = \operatorname{vec}P$ and its Kronecker counterpart $L'L(P'\otimes Q)L' = (P'\otimes Q)L'$, which enable us to discover further properties of the important matrix $L(P'\otimes Q)L'$. Let us now turn away from triangular matrices. An equally important property is $DL\operatorname{vec}A = \operatorname{vec}A$ for symmetric A. Its Kronecker counterpart is $DL(A\otimes A)D = (A\otimes A)D$ for *arbitrary* A, as we shall see shortly, and it enables us to study the matrix $L(A\otimes A)D$.

LEMMA 4.4. *For any (n,n) matrix A,*

(i) $DL(A\otimes A)D = (A\otimes A)D$;

(ii) $$[L(A\otimes A)D]^s = L(A^s\otimes A^s)D, \quad \begin{cases} s=0,1,2,\cdots, \\ s=\cdots,-2,-1 & \text{if } A^{-1} \text{ exists}, \\ s=\frac{1}{2} & \text{if } A^{1/2} \text{ exists}; \end{cases}$$

(iii) *The eigenvalues of* $L(A\otimes A)D$ *are* $\lambda_i\lambda_j$, $i\geqq j$, *when* A *has eigenvalues* λ_i $(i=1\cdots n)$;

(iv) $|L(A\otimes A)D|=|A|^{n+1}$;

(v) $|D'(A\otimes A)D|=2^{(1/2)n(n-1)}|A|^{n+1}$;

If A *is nonsingular*,

(vi) $[D'(A\otimes A)D]^{-1}=LN(A^{-1}\otimes A^{-1})NL'$.

If $AB=BA$, *and* A *and* B *have eigenvalues* λ_i *and* μ_i, $i=1\cdots n$,

(vii) $|D'(A\otimes B)D|=|A||B|\prod_{i>j}(\lambda_i\mu_j+\lambda_j\mu_i)$.

Note that in (vii) we do not require A and B to be symmetric, in contrast to Nel (1978). In applying (vii) one must be careful to note that knowledge of λ_i and μ_j is *not* sufficient in order to compute $\prod_{i>j}(\lambda_i\mu_j+\lambda_j\mu_i)$. In general, it is necessary to carry out the simultaneous reduction of A and B to diagonal form, because the ordering of the eigenvalues is important. A case that can be solved without this reduction is

$$|D'(A^p\otimes A^q)D|=|A|^{p+nq}\prod_{i>j}(\lambda_i^{p-q}+\lambda_j^{p-q}),$$

where p and q are integers (positive, negative or zero).

Similar results hold for the Kronecker sum $I\otimes A+A\otimes I$:

LEMMA 4.5. *For any* (n,n) *matrix* A *with eigenvalues* λ_i $(i=1\cdots n)$,

(i) $DL(I\otimes A+A\otimes I)D=(I\otimes A+A\otimes I)D=2N(I\otimes A)D=2N(A\otimes I)D$;

(ii) *the eigenvalues of* $L(I\otimes A+A\otimes I)D$ *are* $\lambda_i+\lambda_j$, $i\geqq j$;

(iii) $|L(I\otimes A+A\otimes I)D|=2^n|A|\prod_{i>j}(\lambda_i+\lambda_j)$;

(iv) $[L(I\otimes A+A\otimes I)D]^{-1}=L(I\otimes A+A\otimes I)^{-1}D$;

for nonsingular $I\otimes A+A\otimes I$. *The results* (i) *and* (iii) *can be generalized to*

$$\text{(v)}\quad DL\sum_{h=1}^{H}(A^{H-h}\otimes A^{h-1})D=\sum_{h=1}^{H}(A^{H-h}\otimes A^{h-1})D,\qquad H=2,3,\cdots,$$

and

$$\text{(vi)}\quad \left|L\sum_{h=1}^{H}(A^{H-h}\otimes A^{h-1})D\right|=H^n|A|^{H-1}\prod_{i>j}\mu_{ij},\qquad H=2,3,\cdots,$$

where

$$\mu_{ij}=\begin{cases}\dfrac{(\lambda_i^H-\lambda_j^H)}{(\lambda_i-\lambda_j)}, & \text{if } \lambda_i\neq\lambda_j\\ H\lambda_i^{H-1}, & \text{if } \lambda_i=\lambda_j.\end{cases}$$

The next lemma concerns the determinant of the sum or the difference of the matrices $L(A\otimes A)D$ and $L(B\otimes B)D$.

LEMMA 4.6. *Let* A *and* B *be* (n,n) *matrices. Then the determinant*

$$|L(A\otimes A\pm B\otimes B)D|$$

equals

(i) $|A|^{n+1}\prod_{i\geqq j}(1\pm\lambda_i\lambda_j)$, *if* A *is nonsingular and* λ_i, $i=1\cdots n$, *are the eigenvalues of* BA^{-1};

(ii) $\prod_{i\geqq j}(a_{ii}a_{jj}\pm b_{ii}b_{jj})$, *if* $A=(a_{ij})$ *and* $B=(b_{ij})$ *are lower triangular*;

(iii) $\prod_{i\geqq j}(\mu_i\mu_j\pm\theta_i\theta_j)$, *if* $AB=BA$, *where* μ_i *and* θ_i $(i=1\cdots n)$ *denote the eigenvalues of* A *and* B.

Again, knowledge of μ_i and θ_j is, in general, not sufficient to compute (iii). See the remarks under Lemma 4.4(vi).

A final lemma will prove useful in §§5 and 6.

LEMMA 4.7. *Let P be a lower triangular and nonsingular (n, n) matrix, and α a scalar. Let*

(i) $|L[P'\otimes P+\alpha\operatorname{vec}P(\operatorname{vec}P')']L'|=(1+\alpha n)|P|^{n+1}$;

(ii) $[L[P'\otimes P+\alpha\operatorname{vec}P(\operatorname{vec}P')']L')^{-1}=L[P'^{-1}\otimes P^{-1}-\beta\operatorname{vec}P^{-1}(\operatorname{vec}P'^{-1})']L'$,

where $\beta=\alpha/(1+\alpha n)$. Let further A be a symmetric and nonsingular (n, n) matrix. Then,

(iii) $|L[A\otimes A+\alpha\operatorname{vec}A(\operatorname{vec}A)']D|=(1+\alpha n)|A|^{n+1}$;

(iv) $[L[A\otimes A+\alpha\operatorname{vec}A(\operatorname{vec}A)']D]^{-1}=L[A^{-1}\otimes A^{-1}-\beta\operatorname{vec}A^{-1}(\operatorname{vec}A^{-1})']D$.

This ends the theoretical part of this paper.

5. Maximum likelihood estimation of the multivariate normal distribution. We shall now show the usefulness of L and D in a number of applications. Consider a sample of size m from the n-dimensional normal distribution of y with mean μ and positive definite covariance matrix Φ. The maximum likelihood (ML) estimators of μ and Φ are well known, but the derivation of these estimators is often incorrect. The problem is to take properly into account the symmetry conditions on Φ, as has recently been stressed by Richard (1975) and Balestra (1976). More precisely, we should not differentiate the likelihood function with respect to $\operatorname{vec}\Phi$, but with respect to $v(\Phi)$. First we derive the ML estimators of μ and Φ (Lemma 5.1), then the information matrix and asymptotic covariance matrix (Lemma 5.2), and finally we investigate properties of the random vector $v(F)$, an unbiased estimator of $v(\Phi)$.

LEMMA 5.1. *Consider a sample of size m from the n-dimensional normal distribution of y with mean μ and positive definite covariance matrix Φ. The maximum likelihood estimators of μ and Φ are*

$$\hat{\mu}=\left(\frac{1}{m}\right)\sum_i y_i\equiv\bar{y};$$

$$\hat{\Phi}=\left(\frac{1}{m}\right)\sum_i (y_i-\bar{y})(y_i-\bar{y})'.$$

Proof. The loglikelihood function for the sample is

$$\Lambda_m(y;\mu,v(\Phi))=-\frac{1}{2}nm\log 2\pi-\frac{1}{2}m\log|\Phi|-\frac{1}{2}\operatorname{tr}\Phi^{-1}Z,$$

where

$$Z=\sum_{i=1}^{m}(y_i-\mu)(y_i-\mu)'.$$

Using well-known properties of matrix differentials (see (2.10)–(2.13)) and traces (2.4)–(2.5), the first differential of Λ can be written as

$$\begin{aligned}
d\Lambda&=-\frac{1}{2}md\log|\Phi|-\frac{1}{2}\operatorname{tr}(d\Phi^{-1})Z-\frac{1}{2}\operatorname{tr}\Phi^{-1}dZ\\
&=-\frac{1}{2}m\operatorname{tr}\Phi^{-1}d\Phi+\frac{1}{2}\operatorname{tr}\Phi^{-1}(d\Phi)\Phi^{-1}Z\\
&\quad+\frac{1}{2}\operatorname{tr}\Phi^{-1}\left[\sum_i(y_i-\mu)(d\mu)'+(d\mu)\sum_i(y_i-\mu)'\right]\\
&=\frac{1}{2}\operatorname{tr}(d\Phi)\Phi^{-1}(Z-m\Phi)\Phi^{-1}+(d\mu)'\Phi^{-1}\sum_i(y_i-\mu)\\
&=\frac{1}{2}(\operatorname{vec}d\Phi)'(\Phi^{-1}\otimes\Phi^{-1})\operatorname{vec}(Z-m\Phi)+(d\mu)'\Phi^{-1}\sum_i(y_i-\mu)\\
&=\frac{1}{2}(dv(\Phi))'D'(\Phi^{-1}\otimes\Phi^{-1})\operatorname{vec}(Z-m\Phi)+(d\mu)'\Phi^{-1}\sum_i(y_i-\mu).
\end{aligned}$$

Necessary for a maximum is that $d\Lambda=0$ for all $d\mu\neq 0$ and $dv(\Phi)\neq 0$. This gives

$$\Phi^{-1}\sum_i (y_i-\mu)=0,$$

and

$$D'(\Phi^{-1}\otimes\Phi^{-1})\operatorname{vec}(Z-m\Phi)=0.$$

The first condition implies $\hat{\mu}=(1/m)\Sigma y_i\equiv\bar{y}$. The second can be written as

$$D'(\Phi^{-1}\otimes\Phi^{-1})Dv(Z-m\Phi)=0,$$

that is

$$v(Z-m\Phi)=0,$$

since $D'(\Phi^{-1}\otimes\Phi^{-1})D$ is nonsingular. Thus,

$$\hat{\Phi}=(1/m)\hat{Z}=(1/m)\sum_i (y_i-\bar{y})(y_i-\bar{y})'. \quad \square$$

The precision and efficiency of an estimator is usually stated in terms of the information matrix defined by

$$\Psi_m\equiv -E\frac{\partial^2\Lambda_m}{\partial\theta\,\partial\theta'}, \qquad \theta'=(\mu',(v(\Phi))').$$

Its inverse is a lower bound for the covariance matrix of any unbiased estimator of μ and $v(\Phi)$. This is the Cramér-Rao inequality (see, e.g., Rao (1973)). The asymptotic information matrix is defined as

$$\Psi=\lim_{m\to\infty}\frac{1}{m}\Psi_m,$$

and its inverse is the asymptotic covariance matrix of the ML estimator.[1]

LEMMA 5.2. *The information matrix for μ and $v(\Phi)$ is the $(\frac{1}{2}n(n+3),\frac{1}{2}n(n+3))$ matrix*

$$\Psi_m=m\begin{pmatrix}\Phi^{-1} & 0\\ 0 & \frac{1}{2}D'(\Phi^{-1}\otimes\Phi^{-1})D\end{pmatrix};$$

the asymptotic covariance matrix of the ML *estimators $\hat{\mu}$ and $v(\hat{\Phi})$ is*

$$\Psi^{-1}=\begin{pmatrix}\Phi & 0\\ 0 & 2LN(\Phi\otimes\Phi)NL'\end{pmatrix},$$

and the generalized asymptotic variance of $v(\hat{\Phi})$ is

$$|2LN(\Phi\otimes\Phi)NL'|=2^n|\Phi|^{n+1}.$$

Proof. Recall that the first differential of Λ is

$$d\Lambda=(d\mu)'\Phi^{-1}\sum_i (y_i-\mu)+\frac{1}{2}(dv(\Phi))'D'(\Phi^{-1}\otimes\Phi^{-1})\operatorname{vec}(Z-m\Phi).$$

[1]Some authors refer to Ψ_m^{-1} (rather than Ψ^{-1}) as the asymptotic covariance matrix of $\hat{\theta}$.

The second differential is therefore

$$d^2\Lambda=(d\mu)'(d\Phi^{-1})\sum_i(y_i-\mu)-m(d\mu)'\Phi^{-1}(d\mu)$$

$$+\frac{1}{2}(dv(\Phi))'D'(d(\Phi^{-1}\otimes\Phi^{-1}))\mathrm{vec}(Z-m\Phi)$$

$$+\frac{1}{2}(dv(\Phi))'D'(\Phi^{-1}\otimes\Phi^{-1})\mathrm{vec}(dZ-md\Phi).$$

Taking expectations, and observing that $Ey_i=\mu$, $EZ=m\Phi$, and $EdZ=0$, we find

$$-Ed^2\Lambda=m(d\mu)'\Phi^{-1}(d\mu)+\left(\frac{m}{2}\right)(dv(\Phi))'D'(\Phi^{-1}\otimes\Phi^{-1})\mathrm{vec}\,d\Phi$$

$$=m(d\mu)'\Phi^{-1}(d\mu)+\left(\frac{m}{2}\right)(dv(\Phi))'D'(\Phi^{-1}\otimes\Phi^{-1})Ddv(\Phi).$$

The information matrix then follows. From Lemma 4.4 we know that

$$\left[D'(\Phi^{-1}\otimes\Phi^{-1})D\right]^{-1}=LN(\Phi\otimes\Phi)NL',$$

and

$$|D'(\Phi^{-1}\otimes\Phi^{-1})D|=2^{(1/2)n(n-1)}|\Phi|^{-(n+1)}.$$

Hence,

$$\Psi^{-1}=\left(\frac{1}{m}\Psi_m\right)^{-1}=\begin{pmatrix}\Phi & 0\\ 0 & 2LN(\Phi\otimes\Phi)NL'\end{pmatrix},$$

and

$$|2LN(\Phi\otimes\Phi)NL'|=2^{(1/2)n(n+1)}|D'(\Phi^{-1}\otimes\Phi^{-1})D|^{-1}$$

$$=2^{(1/2)n(n+1)}2^{-(1/2)n(n-1)}|\Phi|^{n+1}=2^n|\Phi|^{n+1}. \quad \square$$

The ML estimator $v(\hat{\Phi})$ is *not* an unbiased estimator of $v(\Phi)$. Let us therefore define

$$F\equiv\frac{1}{m-1}\sum_i(y_i-\bar{y})(y_i-\bar{y})'=\frac{m}{m-1}\hat{\Phi}.$$

The following properties of $v(F)$ can then be established.

LEMMA 5.3. *The random vector $v(F)$ is an unbiased estimator of $v(\Phi)$,*

(i) $Ev(F)=v(\Phi)$.

Its covariance matrix is

(ii) $\mathrm{cov}(v(F))=2(LN(\Phi\otimes\Phi)NL')/(m-1)$,

and $v(F)$ is therefore a consistent estimator of $v(\Phi)$. In particular,

(iii) $\mathrm{var}(f_{ij})=(\phi_{ij}^2+\phi_{ii}\phi_{jj})/(m-1), \qquad i\geqq j=1\cdots n.$

Finally, the efficiency of $v(F)$ is

(iv) $\mathrm{eff}(v(F))=[(m-1)/m]^{(1/2)n(n+1)}$.

Proof. We know that

$$m\hat{\Phi}=\sum_i(y_i-\bar{y})(y_i-\bar{y})'=\sum_i y_iy_i'-m\bar{y}\bar{y}'$$

is centrally Wishart distributed $W_n(m-1,\Phi)$, see Rao (1973, p. 537). Therefore, as derived in Magnus and Neudecker (1979, Corollary 4.2),

$$mE\hat{\Phi}=(m-1)\Phi,$$

and

$$\operatorname{cov}(m\operatorname{vec}\hat{\Phi})=(m-1)(I+K)(\Phi\otimes\Phi).$$

Thus, $v(F)=(m/(m-1))v(\hat{\Phi})$ is an unbiased estimator of $v(\Phi)$ and its covariance matrix is

$$\begin{aligned}\operatorname{cov}(v(F))&=\frac{1}{(m-1)^2}\operatorname{cov}(mv(\hat{\Phi}))=\frac{1}{(m-1)^2}\operatorname{cov}(L\operatorname{vec}m\hat{\Phi})\\&=\frac{1}{(m-1)^2}\cdot(m-1)L(I+K)(\Phi\otimes\Phi)L'\\&=\frac{2}{(m-1)}\cdot LN(\Phi\otimes\Phi)L'=\frac{2}{(m-1)}\cdot LN(\Phi\otimes\Phi)NL'.\end{aligned}$$

We see that $\operatorname{cov}(v(F))\to 0$ as $m\to\infty$. This shows that $v(F)$ is a consistent estimator of $v(\Phi)$, given (i). The diagonal elements of $LN(\Phi\otimes\Phi)NL'$ can be derived as follows. Let $i\geqq j$; then, by (2.8), Lemma 3.3 (i), and (2.5),

$$\begin{aligned}u'_{ij}LN(\Phi\otimes\Phi)NL'u_{ij}&=\left(v(E_{ij})\right)'LN(\Phi\otimes\Phi)NL'v(E_{ij})\\&=\tfrac{1}{4}\left(\operatorname{vec}(E_{ij}+E_{ji})\right)'(\Phi\otimes\Phi)\operatorname{vec}(E_{ij}+E_{ji})\\&=\tfrac{1}{2}\left(\operatorname{tr}E_{ij}\Phi E_{ij}\Phi+\operatorname{tr}E_{ij}\Phi E_{ji}\Phi\right)=\tfrac{1}{2}\left(\phi_{ij}^2+\phi_{ii}\phi_{jj}\right).\end{aligned}$$

Thus,

$$\operatorname{var}(f_{ij})=\left(\frac{1}{(m-1)}\right)\left(\phi_{ij}^2+\phi_{ii}\phi_{jj}\right).$$

Finally, the efficiency of $v(F)$ is [see Anderson (1958, p. 57)]

$$\begin{aligned}\operatorname{eff}(v(F))&=\frac{\left|\dfrac{-E\partial^2\Lambda}{\partial v(\Phi)\partial v(\Phi)'}\right|^{-1}}{|\operatorname{cov}(v(F))|}\\&=\left|\frac{m}{2}D'(\Phi^{-1}\otimes\Phi^{-1})D\right|^{-1}\left|\frac{2}{(m-1)}LN(\Phi\otimes\Phi)NL'\right|^{-1}\\&=\left[\frac{(m-1)}{m}\right]^{(1/2)n(n+1)}.\quad\square\end{aligned}$$

Lemmas 5.1 and 5.2 can be straightforwardly generalized by allowing the y_i to have different expectations μ_i. Clearly, it is not possible to estimate all μ_i $(i=1\cdots m)$ and $v(\Phi)$, i.e., $nm+\frac{1}{2}n(n+1)$ parameters, from nm observations. If, however, we assume that the μ_i depend upon a fixed number of parameters $(\theta_1\cdots\theta_K)\equiv\theta'$, and $\hat{\theta}$ is the ML estimator of θ, then the ML estimator of Φ is

$$\hat{\Phi}=\left(\frac{1}{m}\right)\sum_i\left(y_i-\mu_i(\hat{\theta})\right)\left(y_i-\mu_i(\hat{\theta})\right)',$$

and the asymptotic covariance matrix of $v(\hat{\Phi})$ is again

$$\operatorname{as.cov}\left(v(\hat{\Phi})\right)=2LN(\Phi\otimes\Phi)NL'.$$

6. Jacobians. Let the matrix Y be a one-to-one function of a matrix X. The matrix $J=J(Y,X)=(\partial\operatorname{vec}Y/\partial\operatorname{vec}X)'$ is called the Jacobian matrix and its determinant the Jacobian of the transformation of X to Y.

Because the ordering of the variables is arbitrary, the value of the Jacobian can vary in sign, but since only the absolute value matters, this should not worry us. Note that our definition of a Jacobian differs from some textbooks', where $J(Y,X)$ is defined as $|\partial\operatorname{vec}Y/\partial\operatorname{vec}X|^{-1}$.

Consider for example the linear transformation

$$Y=AX,$$

where X and Y are (m,n) matrices, and A is a nonsingular (m,m) matrix. Taking differentials and vecs we have

$$dY=A\,dX,$$

and

$$d\operatorname{vec}Y=(I\otimes A)\,d\operatorname{vec}X,$$

so that

$$|J(Y,X)|=\left|\frac{\partial\operatorname{vec}Y}{\partial\operatorname{vec}X}\right|=|I\otimes A|=|A|^n.$$

The evaluation of Jacobians of transformations involving symmetric or lower triangular matrix arguments is not straightforward, since in this case X contains only $\frac{1}{2}n(n+1)$ "essential" variables. To account for this, a variety of methods have been used, notably differential techniques (Deemer and Olkin (1951) and Olkin (1953)), induction (Jack (1966)), and functional equations induced on the relevant spaces (Olkin and Sampson (1972)). Our approach finds its root in Tracy and Singh (1972) who used modified matrix differentiation results to obtain Jacobians in a simple fashion.

Consider the relation between the $\frac{1}{2}n(n+1)$ variables y_{ij} and the $\frac{1}{2}n(n+1)$ variables x_{ij} given by

$$Y=AXA',$$

where X (and hence Y) is symmetric. Taking differentials and vecs, we have

$$d\operatorname{vec}Y=(A\otimes A)\,d\operatorname{vec}X,$$

and, using the definitions of L and D,

$$dv(Y)=L(A\otimes A)D\,dv(X),$$

so that by Lemma 4.4 (iv)

$$|J(Y,X)|=\left|\frac{\partial v(Y)}{\partial v(X)}\right|=|L(A\otimes A)D|=|A|^{n+1}.$$

See also Deemer and Olkin (1951), Anderson (1958, pp. 156 and 162), Jack (1968), Tracy and Singh (1972), and Olkin and Sampson (1972). Anderson unnecessarily assumes that X has a Wishart distribution or that A is triangular. A more general transformation is

$$Y=AXA'\pm BXB',$$

where X again is symmetric. This yields

$$dv(Y)=L(A\otimes A\pm B\otimes B)D\,dv(X),$$

and the Jacobian matrix is

$$J(Y,X)=L(A\otimes A\pm B\otimes B)D,$$

of which we know the determinant from Lemma 4.6. See Tracy and Singh (1972) for an earlier (wrong) solution in the case $AB=BA$.

We now turn to *nonlinear* transformations involving symmetric matrix arguments. Consider

$$Y=XAX,$$

where A and X are symmetric. Differentiating,

$$dY=(dX)AX+XA(dX),$$

so that

$$d\operatorname{vec}Y=(XA\otimes I+I\otimes XA)d\operatorname{vec}X,$$

and

$$dv(Y)=L(XA\otimes I+I\otimes XA)Ddv(X).$$

Thus, from Lemma 4.5 (iii), the Jacobian is

$$|J(Y,X)|=|L(XA\otimes I+I\otimes XA)D|=2^n|A||X|\prod_{i>j}(\lambda_i+\lambda_j),$$

where λ_i, $i=1\cdots n$, are the eigenvalues of XA. This problem has been studied by Tracy and Singh (1972), but not solved satisfactorily. See also Olkin and Sampson (1972).

The inverse transformation

$$Y=X^{-1},$$

for symmetric X gives

$$dv(Y)=-L(X^{-1}\otimes X^{-1})Ddv(X).$$

Disregarding the minus sign, the Jacobian is (Lemma 4.4 (iv))

$$|J(Y,X)|=|L(X^{-1}\otimes X^{-1})D|=|X|^{-(n+1)}.$$

See Jack (1968), Zellner (1971, pp. 226 and 395), and Olkin and Sampson (1972). Zellner assumes (unnecessarily) that X is positive definite.

More interesting is the transformation, again for $X=X'$,

$$Y=|X|X^{-1}.$$

Totally differentiating yields

$$\begin{aligned} dY&=(d|X|)X^{-1}+|X|dX^{-1}\\ &=|X|(\operatorname{tr}X^{-1}dX)X^{-1}-|X|X^{-1}(dX)X^{-1}, \end{aligned}$$

so that

$$\begin{aligned} d\operatorname{vec}Y&=|X|\left[(\operatorname{vec}X^{-1})(\operatorname{vec}X^{-1})'d\operatorname{vec}X-(X^{-1}\otimes X^{-1})d\operatorname{vec}X\right]\\ &=-|X|\left[X^{-1}\otimes X^{-1}-(\operatorname{vec}X^{-1})(\operatorname{vec}X^{-1})'\right]d\operatorname{vec}X, \end{aligned}$$

and

$$dv(Y)=-|X|L\left[X^{-1}\otimes X^{-1}-(\operatorname{vec}X^{-1})(\operatorname{vec}X^{-1})'\right]Ddv(X).$$

The Jacobian is

$$\begin{aligned} |J(Y,X)|&=|X|^{(1/2)n(n+1)}\left|L\left[X^{-1}\otimes X^{-1}-(\operatorname{vec}X^{-1})(\operatorname{vec}X^{-1})'\right]D\right|\\ &=|X|^{(1/2)n(n+1)}(1-n)|X|^{-(n+1)} \qquad \text{(by Lemma 4.7 (iii))}\\ &=-(n-1)|X|^{(1/2)(n+1)(n-2)}. \end{aligned}$$

See Deemer and Olkin (1951) for a solution along completely different lines, assuming X to be positive definite rather than only symmetric.

As a final example of the usefulness of L and D in evaluating Jacobians of transformations with symmetric matrix arguments, consider

$$Y = X^p, \qquad p = 2, 3, \cdots.$$

Upon differentiating we find

$$\begin{aligned} dY &= (dX)X^{p-1} + X(dX)X^{p-2} + \cdots + X^{p-1}(dX) \\ &= \sum_{h=1}^{p} X^{h-1}(dX)X^{p-h}, \end{aligned}$$

which gives

$$d\,\mathrm{vec}\,Y = \sum_{h=1}^{p} (X^{p-h} \otimes X^{h-1})\,d\,\mathrm{vec}\,X,$$

and

$$dv(Y) = L \sum_{h=1}^{p} (X^{p-h} \otimes X^{h-1}) D\,dv(X),$$

so that the Jacobian matrix is

$$J(Y, X) = L \sum_{h=1}^{p} (X^{p-h} \otimes X^{h-1}) D,$$

the determinant of which is given in Lemma 4.5 (vi).

Summarizing, we have considered six relations between the $\frac{1}{2}n(n+1)$ variables of a symmetric matrix Y and the $\frac{1}{2}n(n+1)$ variables of a symmetric matrix X. The results are given in Table 6.1.

Let us now investigate transformations involving *lower triangular* matrix arguments. Consider the relation between lower triangular Y and lower triangular X given by

$$Y = PXQ,$$

where P and Q are also lower triangular. We find

$$d\,\mathrm{vec}\,Y = (Q' \otimes P)\,d\,\mathrm{vec}\,X,$$

and thus

$$dv(Y) = L(Q' \otimes P)L'dv(X).$$

Hence, the Jacobian is

$$|J(Y, X)| = |L(Q' \otimes P)L'| = \prod_i p_{ii}^{i} q_{ii}^{n-i+1} \qquad \text{(Lemma 4.3 (i))}.$$

This problem has been solved by Olkin and Sampson (1972), Deemer and Olkin (1951) for $Q = I$, and Olkin (1953) for $P = I$.

More general is the transformation

$$Y = PXQ + RXS,$$

with lower triangular P, Q, R, S. This leads to

$$dv(Y) = L(Q' \otimes P + S' \otimes R)L'dv(X),$$

and the Jacobian follows from Lemma 4.3 (ii).

TABLE 6.1
Jacobians of transformations with symmetric matrix arguments

	Transformation	Jacobian $\lvert J(Y,X)\rvert$	Conditions, particularities
(i)	$Y=AXA'$	$\lvert A\rvert^{n+1}$	$\lvert A\rvert\neq 0$
(ii)	$Y=AXA'+BXB'$	$\lvert A\rvert^{n+1}\prod_{i\geq j}(1\pm\lambda_i\lambda_j)$	$\lvert A\rvert\neq 0$, $\lambda_i\ (i=1\cdots n)$ eigenvalues of BA^{-1}
		$\prod_{i\geq j}(a_{ii}a_{jj}\pm b_{ii}b_{jj})$	$A=(a_{ij})$ and $B=(b_{ij})$ lower triangular
		$\prod_{i\geq j}(\mu_i\mu_j\pm\theta_i\theta_j)$*	$AB=BA$, μ_i and $\theta_i\ (i=1\cdots n)$ eigenvalues of A and B
(iii)	$Y=XAX$	$2^n\lvert A\rvert\lvert X\rvert\prod_{i>j}(\lambda_i+\lambda_j)$	$A=A'$, $\lambda_i\ (i=1\cdots n)$ eigenvalues of XA
(iv)	$Y=X^{-1}$	$\lvert X\rvert^{-(n+1)}$	$\lvert X\rvert\neq 0$
(v)	$Y=\lvert X\rvert X^{-1}$	$(n-1)\lvert X\rvert^{(1/2)(n+1)(n-2)}$	$\lvert X\rvert\neq 0$
(vi)	$Y=X^p\ (p=2,3,\cdots)$	$p^n\lvert X\rvert^{p-1}\prod_{i>j}\mu_{ij}$	$\mu_{ij}=\begin{cases}(\lambda_i^p-\lambda_j^p)/(\lambda_i-\lambda_j), & \text{if } \lambda_i\neq\lambda_j,\\ p\lambda_i^{p-1}, & \text{if } \lambda_i=\lambda_j,\end{cases}$ where $\lambda_i\ (i=1\cdots n)$ are eigenvalues of X

*See the remarks under Lemmas 4.4 (vi) and 4.6 (iii).

Next, we consider the relation between symmetric Y and lower triangular X given by

$$Y=B'XA+A'X'B.$$

Using the same technique, we have

$$\begin{aligned}d\operatorname{vec}Y&=(A'\otimes B')d\operatorname{vec}X+(B'\otimes A')d\operatorname{vec}X'\\&=\left[A'\otimes B'+(B'\otimes A')K\right]d\operatorname{vec}X\qquad\text{(Definition 2.1a)},\end{aligned}$$

and

$$dv(Y)=L\left[A'\otimes B'+(B'\otimes A')K\right]L'dv(X).$$

The Jacobian is

$$\begin{aligned}\lvert J(Y,X)\rvert&=\lvert L[A'\otimes B'+(B'\otimes A')K]L'\rvert=\lvert L[(A'\otimes B')+K(A'\otimes B')]L'\rvert\\&=\lvert 2LN(A'\otimes B')L'\rvert=2^{(1/2)n(n+1)}\lvert L(A\otimes B)NL'\rvert\\&=2^{(1/2)n(n+1)}\lvert L(A\otimes B)DLNL'\rvert=2^{(1/2)n(n+1)}2^{-(1/2)n(n-1)}\lvert L(A\otimes B)D\rvert\\&=2^n\lvert L(A\otimes B)D\rvert,\end{aligned}$$

by the definition of N and Lemmas 2.1 (ii), 3.5 (ii), and 3.4 (i). The determinant $\lvert L(A\otimes B)D\rvert$ can of course be evaluated for each specific A and B. In particular, if $A=P$ and $B=Q'R$, or $A=PQ'$ and $B=R'$, where P, Q, and R are lower triangular, we can express this determinant in terms of the diagonal elements of P, Q, and R, by Lemma 4.3 (iii)–(iv). Special cases have been solved by Deemer and Olkin (1951) ($A=P'$ and $B=I$) and by Olkin (1953) ($A=I$ and $B=P$).

Turning now to *nonlinear* transformations involving lower triangular matrix arguments, we first consider the relation

$$Y=XPX,$$

with lower triangular P. We find

$$dY=(dX)PX+XP(dX),$$

and

$$dv(Y)=L(X'P'\otimes I+I\otimes XP)L'dv(X).$$

By Lemma 4.3 (ii) the Jacobian is

$$|J(Y,X)|=|L(X'P'\otimes I+I\otimes XP)L'|=\prod_{i\geqq j}(p_{jj}x_{jj}+p_{ii}x_{ii})$$

$$=2^n|P||X|\prod_{i>j}(p_{ii}x_{ii}+p_{jj}x_{jj}).$$

The next transformation is between a symmetric Y and lower triangular X,

$$Y=X'AX+XBX',$$

where $A=A'$ and $B=B'$. Proceeding as before we find

$$d\operatorname{vec}Y=(I\otimes X'A+XB\otimes I)d\operatorname{vec}X+(X'A\otimes I+I\otimes XB)d\operatorname{vec}X'$$

$$=2N(I\otimes X'A+XB\otimes I)d\operatorname{vec}X,$$

so that

$$dv(Y)=2LN(I\otimes X'A+XB\otimes I)L'dv(X).$$

The Jacobian is thus

$$|J(Y,X)|=2^{(1/2)n(n+1)}|L(I\otimes AX+BX'\otimes I)NL'|$$

$$=2^n|L(I\otimes AX+BX'\otimes I)D|.$$

Special cases are the transformations $Y=XX'$ ($A=0$, $B=I$) and $Y=X'X$ ($A=I$, $B=0$), for which the Jacobians can be expressed in terms of the diagonal elements of X by Lemma 4.3 (iii)–(iv). See Deemer and Olkin (1951), Olkin (1953), Jack (1966), Olkin and Sampson (1972), and Zellner (1971, p. 392).

The Jacobians of the transformations $Y=X^{-1}$, $Y=|X|X^{-1}$, and $Y=X^p$, $p=2,3,\cdots$, for lower triangular X and Y can be determined in a fashion very similar to their symmetric counterparts. For $Y=X^{-1}$ we find

$$dv(Y)=-L(X'^{-1}\otimes X^{-1})L'dv(X).$$

For $Y=|X|X^{-1}$,

$$dv(Y)=-|X|L\left[X'^{-1}\otimes X^{-1}-\operatorname{vec}X^{-1}(\operatorname{vec}X'^{-1})'\right]L'dv(X),$$

and for $Y=X^p$,

$$dv(Y)=L\sum_{h=1}^{p}\left[(X')^{p-h}\otimes X^{h-1}\right]L'dv(X).$$

The Jacobians of the three transformations are easily recognized as determinants which have been studied in §4 (Lemmas 4.3 (i), 4.7 (i) and 4.3 (vi)).

The above discussion about Jacobians of transformations with lower triangular matrix arguments is summarized in Table 6.2.

TABLE 6.2

Jacobians of transformations with lower triangular matrix arguments

	Transformation	Jacobian $\|J(Y,X)\|$	Conditions, particularities
(i)	$Y=PXQ$	$\prod_i p_{ii}^i q_{ii}^{n-i+1}$	P,Q lower triangular
(ii)	$Y=PXQ+RXS$	$\prod_{i\geq j}(p_{ii}q_{jj}+r_{ii}s_{jj})$	P,Q,R,S lower triangular
(iii)	$Y=B'XA+A'X'B$	$2^n\|L(A\otimes B)D\|$	
(iiia)	$Y=R'QXP+P'X'Q'R$	$2^n\prod_i(q_{ii}r_{ii})^i p_{ii}^{n-i+1}$	P,Q,R lower triangular
(iiib)	$Y=RXPQ'+QP'X'R'$	$2^n\prod_i(p_{ii}q_{ii})^{n-i+1}r_{ii}^i$	P,Q,R lower triangular
(iv)	$Y=XPX$	$2^n\|P\|\|X\|\prod_{i>j}(p_{ii}x_{ii}+p_{jj}x_{jj})$	P lower triangular
(v)	$Y=X'AX+XBX'$	$2^n\|L(I\otimes AX+BX'\otimes I)D\|$	$A=A'$, $B=B'$
(va)	$Y=XX'$	$2^n\prod_i x_{ii}^{n-i+1}$	
(vb)	$Y=X'X$	$2^n\prod_i x_{ii}^i$	
(vi)	$Y=X^{-1}$	$\|X\|^{-(n+1)}$	$\|X\|\neq 0$
(vii)	$Y=\|X\|X^{-1}$	$(n-1)\|X\|^{(1/2)(n+1)(n-2)}$	$\|X\|\neq 0$
(viii)	$Y=X^p(p=2,3,\cdots)$	$p^n\|X\|^{p-1}\prod_{i>j}\mu_{ij}$	$\mu_{ij}=\begin{cases}(x_{ii}^p-x_{jj}^p)/(x_{ii}-x_{jj}) & \text{if } x_{ii}\neq x_{jj},\\ px_{ii}^{p-1} & \text{if } x_{ii}=x_{jj},\end{cases}$ where x_{ii} $(i=1\cdots n)$ are the diagonal elements of X

7. **Matrix equations.** A third area where we can demonstrate the usefulness of the matrices L and D is the solution of matrix equations. Suppose we are given a matrix equation $Y=F(X)$, where we know a priori that X is symmetric (or triangular). We wish to solve X in terms of Y. If Y is a one-to-one function of X, as in the preceding section on Jacobians, then $X=F^{-1}(Y)$ is the unique solution. If, however, Y is not in one-to-one correspondence with X, we have to restrict the solution space of X to symmetric (or triangular) matrices. In other words, we should not solve for X, but for $v(X)$. An example may clarify this approach.

LEMMA 7.1. *The vector equation*

$$Q\,\mathrm{vec}X=\mathrm{vec}A,$$

where Q and A are (n^2,n^2) and (n,n) matrices respectively, and X is known to be symmetric, has a solution for X if and only if

$$QD(QD)^+\mathrm{vec}A=\mathrm{vec}A,$$

in which case the general solution is

$$\mathrm{vec}X=D(QD)^+\mathrm{vec}A+D\big[I-(QD)^+QD\big]\mathrm{vec}P,$$

where $\mathrm{vec}P$ *is an arbitrary* $\frac{1}{2}n(n+1)$*-vector and* $(QD)^+$ *denotes the Moore-Penrose inverse of* QD.

Proof. Since X is symmetric, we have $\mathrm{vec}X=Dv(X)$ and thus $QDv(X)=\mathrm{vec}A$. The consistency and solution of this system follow from Penrose (1955, p. 409). Thus, if a solution exists, it has the form

$$v(X)=(QD)^{+}\mathrm{vec}A+\left[I-(QD)^{+}QD\right]\mathrm{vec}P,$$

for arbitrary P. Premultiplication with D gives the desired result. □

This problem has also been studied by Vetter (1975, p. 187), but not solved satisfactorily. If Q is nonsingular, a solution exists if and only if $(I-K)Q^{-1}\mathrm{vec}A=0$. The (unique) solution then takes the form $\mathrm{vec}X=Q^{-1}\mathrm{vec}A$. If LQD is nonsingular, we may write the solution, if it exists, as $\mathrm{vec}X=D(LQD)^{-1}L\mathrm{vec}A$. This is Vetter's solution. He assumes nonsingularity of Q and of LQD (neither of which implies the other!) and also tacitly the existence of a solution for $\mathrm{vec}X$. Note that if we know X to be lower triangular rather than symmetric, the solution is obtained by replacing D with L'.

As a final example, let us consider a problem which arises in dynamic econometric models. It concerns the equilibrium covariance matrix. We want to find the matrix of partial derivatives of S with respect to A for $S=ASA'+V$, with symmetric V and S, when all eigenvalues of A are less than 1 in absolute value. This problem was first studied by Conlisk (1969) who derived $\partial\,\mathrm{vec}S/\partial a_{ij}$ for each element of A separately. Neudecker (1969) gave a compact expression for $\partial\,\mathrm{vec}S/\partial\,\mathrm{vec}A$. His derivation is wrong, but the result is correct.

LEMMA 7.2. *Consider the matrix equation*

$$S=ASA'+V,$$

when S and V are symmetric (n,n) matrices, and all eigenvalues of A are smaller than 1 *in absolute value. The partial derivatives of S with respect to A can be expressed as*

$$\left(\frac{\partial\,\mathrm{vec}S}{\partial\,\mathrm{vec}A}\right)'=2N(I\otimes I-A\otimes A)^{-1}(AS\otimes I).$$

The partial derivatives of the distinct elements of S with respect to A can be expressed as

$$\left(\frac{\partial v(S)}{\partial\,\mathrm{vec}A}\right)'=2LN(I\otimes I-A\otimes A)^{-1}(AS\otimes I).$$

Proof. We take differentials and vecs:

$$dS=A(dS)A'+(dA)SA'+AS(dA)'+dV,$$

$$d\,\mathrm{vec}S=(A\otimes A)d\,\mathrm{vec}S+(AS\otimes I)d\,\mathrm{vec}A+(I\otimes AS)d\,\mathrm{vec}A'+d\,\mathrm{vec}V.$$

Using Lemma 2.1 (ii), and the definitions of K and N, we have

$$\begin{aligned}(I\otimes I-A\otimes A)d\,\mathrm{vec}S&=\left[AS\otimes I+(I\otimes AS)K\right]d\,\mathrm{vec}A+d\,\mathrm{vec}V\\&=\left[AS\otimes I+K(AS\otimes I)\right]d\,\mathrm{vec}A+d\,\mathrm{vec}V\\&=2N(AS\otimes I)d\,\mathrm{vec}A+d\,\mathrm{vec}V.\end{aligned}$$

Since the eigenvalues of A are smaller than 1 in absolute value, the matrix $I\otimes I-A\otimes A$ is nonsingular, and

$$(I\otimes I-A\otimes A)^{-1}=\sum_{h=0}^{\infty}(A^{h}\otimes A^{h}).$$

Thus, by Lemma 2.1 (v),

$$d\mathrm{vec}S = 2\sum_h (A^h \otimes A^h) N (AS \otimes I) d\mathrm{vec}A + (I\otimes I - A\otimes A)^{-1} d\mathrm{vec}V$$

$$= 2N\sum_h (A^h \otimes A^h)(AS \otimes I) d\mathrm{vec}A + (I\otimes I - A\otimes A)^{-1} d\mathrm{vec}V$$

$$= 2N(I\otimes I - A\otimes A)^{-1}(AS\otimes I) d\mathrm{vec}A + (I\otimes I - A\otimes A)^{-1} d\mathrm{vec}V,$$

and $$dv(S) = Ld\mathrm{vec}S = 2LN(I\otimes I - A\otimes A)^{-1}(AS\otimes I)d\mathrm{vec}A$$

$$+ L(I\otimes I - A\otimes A)^{-1} d\mathrm{vec}V. \qquad \square$$

Appendix: proofs of lemmas in §4.

Proof of Lemma 4.1. Let A and B be (n, n) matrices. We shall write $L(A\otimes B)L'$ in terms of unit vectors, using the explicit definition of L and (2.5).

$$L(A\otimes B)L' = \sum_{i\geqq j} u_{ij}(\mathrm{vec}E_{ij})'(A\otimes B)\sum_{s\geqq t}(\mathrm{vec}E_{st})u'_{st}$$

$$= \sum_{i\geqq j}\sum_{s\geqq t}(\mathrm{vec}E_{ij})'(A\otimes B)(\mathrm{vec}E_{st})u_{ij}u'_{st}$$

$$= \sum_{i\geqq j}\sum_{s\geqq t}\mathrm{tr}(E_{ji}BE_{st}A')u_{ij}u'_{st} = \sum_{i\geqq j}\sum_{s\geqq t} a_{jt}b_{is}u_{ij}u'_{st}.$$

If $A = \Lambda$, $B = M$, and δ_{ij} denotes the Kronecker delta symbol ($\delta_{ij} = 0$ if $i\neq j$, $\delta_{ii} = 1$), we find

$$L(\Lambda\otimes M)L' = \sum_{i\geqq j}\sum_{s\geqq t}\delta_{jt}\delta_{is}\lambda_j\mu_i u_{ij}u'_{st} = \sum_{i\geqq j}\lambda_j\mu_i u_{ij}u'_{ij},$$

which is a diagonal matrix (since $u_{ij}u'_{ij}$ is diagonal) with elements $\mu_i\lambda_j$, $i\geqq j$, and determinant $\Pi_{i\geqq j}\mu_i\lambda_j = \Pi_i \mu_i^i \lambda_i^{n-i+1}$.

If $A = P$ and $B = Q$, we have $L(P\otimes Q)L' = \Sigma_{i\geqq j}\Sigma_{s\geqq t}\, p_{jt}q_{is}u_{ij}u'_{st}$. Because P and Q are lower triangular, we may restrict the summation to $i\geqq j\geqq t$, $i\geqq s\geqq t$. This implies that the matrix $u_{ij}u'_{st}$ is lower triangular. Hence, $L(P\otimes Q)L'$ is lower triangular. By putting $s = i$ and $t = j$, we find that its diagonal elements are $q_{ii}p_{jj}$, $i\geqq j$, and its determinant is $\Pi_{i\geqq j}\, q_{ii}p_{jj} = \Pi_i q_{ii}^i p_{ii}^{n-i+1}$.

Similarly, we can express $L(A\otimes B)D$ in terms of unit vectors, using the explicit definitions of L and D. One verifies that

$$L(A\otimes B)D = L(A\otimes B)L' + \Gamma(A, B),$$

where

$$\Gamma(A, B) = \sum_{i\geqq j}\sum_{s>t} a_{js}b_{it}u_{ij}u'_{st}.$$

Consider the matrix $\Gamma(A, B)$. It is easy to see that $\Gamma(\Lambda, M) = 0$. If $A = P$ and $B = Q$, we may restrict the summation to $i\geqq j\geqq s>t$, so that $\Gamma(P, Q)$ is *strictly* lower triangular. If $A = P'$ and $B = Q'$, we may restrict the summation to $s>t\geqq i\geqq j$, so that $\Gamma(P', Q')$ is *strictly* upper triangular. The properties of $L(\Lambda\otimes M)D$, $L(P\otimes Q)D$, and $L(P'\otimes Q')D$ then follow from the properties of $L(\Lambda\otimes M)L'$ and $L(P\otimes Q)L'$. $\square$

Proof of Lemma 4.2. Let P and Q be lower triangular and $v(X)$ arbitrary. Remembering that $L'v(X)=\text{vec}\bar{X}$ (Lemma 3.3 (i)), we have

$$L'L(P'\otimes Q)L'v(X)=L'L(P'\otimes Q)\text{vec}\bar{X}=L'L\text{vec}Q\bar{X}P$$
$$=\text{vec}Q\bar{X}P=(P'\otimes Q)\text{vec}\bar{X}=(P'\otimes Q)L'v(X).$$

Thus,

$$L'L(P'\otimes Q)L'=(P'\otimes Q)L'.$$

Property (ii) follows from repeated application of (i). Further, $D'(P'\otimes Q)L'=D'L'L(P'\otimes Q)L'=L(P'\otimes Q)L'$, since $LD=I$ (Lemma 3.5(i)). Let us now determine the eigenvalues of $L(P'\otimes Q)L'$. We will need the following result.

Result A.1. Let P be a lower triangular matrix with *distinct* diagonal elements. Then there exists a lower triangular matrix S with ones on the diagonal such that $S^{-1}PS=\text{dg}(P)$.

Proof. Consider the $\frac{1}{2}n(n-1)$ equations in $\frac{1}{2}n(n-1)$ unknowns $(s_{ij}, i>j)$ given by $PS=S\text{dg}(P)$. This gives $p_{ij}+\Sigma_{h=j+1}^{i}p_{ih}s_{hj}=s_{ij}p_{jj}$, $i>j$, from which we can sequentially solve for $s_{j+1,j}$ $(j=1\cdots n-1)$, $s_{j+2,j}$ $(j=1\cdots n-2),\cdots,s_{n1}$. □

Assume that both P and Q have distinct diagonal elements. Then, by Result A.1, there exist lower triangular matrices S and T with ones on the diagonal such that

$$S^{-1}PS=\text{dg}(P) \quad\text{and}\quad T^{-1}QT=\text{dg}(Q).$$

By repeated application of (i) we see that

$$L(S'\otimes T^{-1})L'L(P'\otimes Q)L'L(S'^{-1}\otimes T)L'$$
$$=L(S'P'S'^{-1}\otimes T^{-1}QT)L'=L(\text{dg}(P)\otimes\text{dg}(Q))L',$$

and

$$L(S'\otimes T^{-1})L'L(S'^{-1}\otimes T)L'=LL'=I.$$

From Lemma 4.1 we know that $L(\text{dg}(P)\otimes\text{dg}(Q))L'$ is a diagonal matrix with elements $q_{ii}p_{jj}$, $i\geqq j$. These, therefore, are the eigenvalues of $L(P'\otimes Q)L'$.

If not all diagonal elements of P and Q are distinct, we can obtain (iii) by way of a limiting relation, starting with $P+\Delta$ and $Q+\Delta$, where Δ is a diagonal matrix with δ^h as its hth diagonal element. If δ is sufficiently small, $P+\Delta$ and $Q+\Delta$ will each have distinct diagonal elements. Hence the eigenvalues of $L[(P+\Delta)'\otimes(Q+\Delta)]L'$ are $(q_{ii}+\delta^i)(p_{jj}+\delta^j)$, $i\geqq j$. Letting $\delta\to 0$, we find the desired result. □

Proof of Lemma 4.3. Using Lemma 4.2, we have

$$|L(PQ'\otimes R'S)L'|=|L(P\otimes R')(Q'\otimes S)L'|=|L(P\otimes R')L'L(Q'\otimes S)L'|$$
$$=|L(P'\otimes R)L'||L(Q'\otimes S)L'|=\Big(\prod_{i\geqq j}r_{ii}p_{jj}\Big)\Big(\prod_{i\geqq j}s_{ii}q_{jj}\Big)$$
$$=\prod_i (r_{ii}s_{ii})^i(p_{ii}q_{ii})^{n-i+1}.$$

To prove (ii) we first assume that P and Q are nonsingular. Then, by Lemma 4.2 and $LL'=I$,

$$L(P'\otimes Q+R'\otimes S)L'=L(I\otimes I+R'P'^{-1}\otimes SQ^{-1})(P'\otimes Q)L'$$
$$=L(I\otimes I+R'P'^{-1}\otimes SQ^{-1})L'L(P'\otimes Q)L'$$
$$=(I+L(R'P'^{-1}\otimes SQ^{-1})L')(L(P'\otimes Q)L').$$

Hence,

$$|L(P'\otimes Q+R'\otimes S)L'|=\prod_{i\geqq j}\left(1+\frac{s_{ii}}{q_{ii}}\frac{r_{jj}}{p_{jj}}\right)\prod_{i\geqq j}(q_{ii}p_{jj})=\prod_{i\geqq j}(q_{ii}p_{jj}+s_{ii}r_{jj}).$$

If P or Q is singular, we obtain (ii) starting with $P+\delta I$ or $Q+\delta I$, where δ is small and $P+\delta I$ or $Q+\delta I$ is nonsingular. To prove (iii) and (iv) we use Lemmas 4.1 (iii) and 4.2 (i) and (iii):

$$|L(P\otimes Q'R)D|=|L(P\otimes Q')(I\otimes R)D|=|L(P\otimes Q')L'L(I\otimes R)D|$$

$$=|L(P'\otimes Q)L'||L(I\otimes R)D|=\Pi_i q_{ii}^i p_{ii}^{n-i+1}\Pi_i r_{ii}^i=\Pi_i(q_{ii}r_{ii})^i p_{ii}^{n-i+1},$$

$$|L(PQ'\otimes R')D|=|L(P\otimes R')(Q'\otimes I)D|=|L(P\otimes R')L'L(Q'\otimes I)D|$$

$$=|L(P'\otimes R)L'||L(Q'\otimes I)D|=\Pi_i r_{ii}^i p_{ii}^{n-i+1}\Pi_i q_{ii}^{n-i+1}=\Pi_i r_{ii}^i(p_{ii}q_{ii})^{n-i+1}.$$

For nonsingular P, Q, R, S, we again use Lemma 4.2 (i) to prove (v) as follows.

$$L(PQ'\otimes R'S)L'L(Q'^{-1}\otimes S^{-1})L'L(P^{-1}\otimes R'^{-1})L'$$

$$=L(P\otimes R')L'L(P^{-1}\otimes R'^{-1})L'=LL'=I.$$

Let us now show (vi). Assume that P has distinct diagonal elements. Then there exists a lower triangular matrix S with ones on the diagonal such that $S^{-1}PS=\Lambda$, with $\Lambda=\mathrm{dg}(P)$ to simplify notation (see Result A.1). Thus,

$$L\sum_{h=1}^{H}\left[(P')^{H-h}\otimes P^{h-1}\right]L'=\sum_h L\left[S'^{-1}\Lambda^{H-h}S'\otimes S\Lambda^{h-1}S^{-1}\right]L'$$

$$=\sum_h L(S'^{-1}\otimes S)L'L(\Lambda^{H-h}\otimes\Lambda^{h-1})L'L(S'\otimes S^{-1})L'$$

$$=(L(S'^{-1}\otimes S)L')\sum_h(L(\Lambda^{H-h}\otimes\Lambda^{h-1})L')(L(S'\otimes S^{-1})L').$$

Because

$$(L(S'^{-1}\otimes S)L')^{-1}=L(S'\otimes S^{-1})L',$$

we have

$$\left|L\sum_{h=1}^{H}\left[(P')^{H-h}\otimes P^{h-1}\right]L'\right|=\left|\sum_h L(\Lambda^{H-h}\otimes\Lambda^{h-1})L'\right|.$$

Now, from Lemma 4.1 we know that $L(\Lambda^{H-h}\otimes\Lambda^{h-1})L'$ is a diagonal matrix with diagonal elements $\lambda_i^{h-1}\lambda_j^{H-h}$, $i\geqq j$. Hence $\Sigma_{h=1}^H L(\Lambda^{H-h}\otimes\Lambda^{h-1})L'$ is diagonal with elements $\Sigma_h\lambda_i^{h-1}\lambda_j^{H-h}$, $i\geqq j$, and determinant

$$\left|\sum_h L(\Lambda^{H-h}\otimes\Lambda^{h-1})L'\right|=\prod_{i\geqq j}\left(\sum_{h=1}^{H}\lambda_i^{h-1}\lambda_j^{H-h}\right)=\prod_i(H\lambda_i^{H-1})\prod_{i>j}\left(\sum_{h=1}^{H}\lambda_i^{h-1}\lambda_j^{H-h}\right)$$

$$=H^n\left(\prod_i\lambda_i\right)^{H-1}\prod_{i>j}\left(\sum_{h=1}^{H}\lambda_i^{h-1}\lambda_j^{H-h}\right)=H^n|P|^{H-1}\prod_{i>j}\mu_{ij},$$

where $\mu_{ij}=\Sigma_{h=1}^H p_{ii}^{h-1}p_{jj}^{H-h}$ (since $\lambda_i=p_{ii}$)$=(p_{ii}^H-p_{jj}^H)/(p_{ii}-p_{jj})$. The case where not all diagonal elements of P are distinct, say $p_{ii}=p_{jj}$, can be considered to be a limiting case of the situation where p_{jj} approaches p_{ii}. Taking the limit as $p_{jj}\to p_{ii}$, we find $\mu_{ij}=Hp_{ii}^{H-1}$. □

Proof of Lemma 4.4. Using the properties $DLN=N$, $D=ND$, and $(A\otimes A)N=N(A\otimes A)$—see Lemmas 3.5 (ii), 3.6 (ii), and 2.1 (v)—we have

$$DL(A\otimes A)D=DL(A\otimes A)ND=DLN(A\otimes A)D=N(A\otimes A)D$$
$$=(A\otimes A)ND=(A\otimes A)D.$$

This proves (i). By repeated application of (i) we find (ii). To prove (iii) we note that $L(A\otimes A)D$ and $DL(A\otimes A)$ have the same set of eigenvalues apart from $\frac{1}{2}n(n-1)$ zeroes which belong to the latter matrix. Let A have eigenvalues λ_i and eigenvectors x_i; then

$$DL(A\otimes A)(x_i\otimes x_j+x_j\otimes x_i)=DL(Ax_i\otimes Ax_j+Ax_j\otimes Ax_i)$$
$$=\lambda_i\lambda_j DL(x_i\otimes x_j+x_j\otimes x_i)=\lambda_i\lambda_j DL\text{vec}(x_jx_i'+x_ix_j')$$
$$=\lambda_i\lambda_j\text{vec}(x_jx_i'+x_ix_j') \text{ (by the implicit definition of } D)$$
$$=\lambda_i\lambda_j(x_i\otimes x_j+x_j\otimes x_i).$$

Hence, $DL(A\otimes A)$ has eigenvalues $\lambda_i\lambda_j$, $i\geqq j$, plus $\frac{1}{2}n(n-1)$ zeroes, and $L(A\otimes A)D$ has eigenvalues $\lambda_i\lambda_j$, $i\geqq j$. Its determinant is

$$|L(A\otimes A)D|=\prod_{i\geqq j}\lambda_i\lambda_j=\prod_i\lambda_i^{n+1}=|A|^{n+1}.$$

Let us now prove (v) and (vi). Since $D=NL'(LNL')^{-1}$ (Lemma 3.5 (iii)), and again using Lemmas 3.6 (ii) and 2.1 (v), we can write

$$D'(A\otimes A)D=(LNL')^{-1}LN(A\otimes A)D=(LNL')^{-1}L(A\otimes A)D.$$

The properties of $D'(A\otimes A)D$ thus follow from the properties of LNL' (Lemma 3.4) and $L(A\otimes A)D$ (this lemma).

To prove (vii) we first assume that A has distinct eigenvalues. In that case there exists a matrix T such that $T^{-1}AT=\Lambda$, where Λ is a diagonal matrix containing the eigenvalues of A. From $AB=BA$ we have $T\Lambda T^{-1}B=BT\Lambda T^{-1}$, or $\Lambda M=M\Lambda$, where $M=T^{-1}BT$. Since all λ's are distinct by assumption, M is diagonal. Hence it contains the eigenvalues of B. We may then write

$$D'(A\otimes B)D=D'(T\Lambda T^{-1}\otimes TMT^{-1})D=D'(T\otimes T)(\Lambda\otimes M)(T^{-1}\otimes T^{-1})D$$
$$=D'(T\otimes T)L'D'(\Lambda\otimes M)DL(T^{-1}\otimes T^{-1})D,$$

by (i). Hence, using (iv), the explicit definition of N, and Lemmas 3.5 (iii), 2.1 (ii), 3.4 (i), and 3.6 (ii),

$$|D'(A\otimes B)D|=|D'(\Lambda\otimes M)D|=|(LNL')^{-1}LN(\Lambda\otimes M)D|$$
$$=|LNL'|^{-1}|\frac{1}{2}L(I+K)(\Lambda\otimes M)D|$$
$$=|LNL'|^{-1}2^{-(1/2)n(n+1)}|L(\Lambda\otimes M)D+L(M\otimes\Lambda)KD|$$
$$=2^{-n}|L(\Lambda\otimes M+M\otimes\Lambda)D|.$$

From Lemma 4.1 we know that $L(\Lambda\otimes M)D$ and $L(M\otimes\Lambda)D$ are diagonal matrices with elements $\mu_i\lambda_j$ and $\lambda_i\mu_j$, $i\geqq j$. The determinant of their sum is $\prod_{i\geqq j}(\mu_i\lambda_j+\lambda_i\mu_j)$, and thus

$$|D'(A\otimes B)D|=2^{-n}\prod_{i\geq j}(\mu_i\lambda_j+\lambda_i\mu_j)=2^{-n}\prod_{i}(2\lambda_i\mu_i)\prod_{i>j}(\mu_i\lambda_j+\lambda_i\mu_j)$$
$$=|A||B|\prod_{i>j}(\mu_i\lambda_j+\lambda_i\mu_j).$$

If A has multiple eigenvalues, say $\lambda_i=\lambda_j$, we consider this as a limiting case of the situation where λ_j approaches λ_i. Taking the limit as $\lambda_j\to\lambda_i$ the result follows. □

Proof of Lemma 4.5. From Lemma 2.1 (vi) and the properties $ND=D$ and $DLN=N$, follows (i). The proof of (ii) is similar to that of Lemma 4.4 (iii). Property (iii) follows from (ii). Property (iv) follows from (i) since $LD=I$. We find (v) by repeated application of $DL(I\otimes A+A\otimes I)D=(I\otimes A+A\otimes I)D$, and Lemma 4.4 (i). Let us prove (vi). We proceed as in the proof of Lemma 4.3 (vi). If A has distinct eigenvalues (or if $A=A'$), there exists a nonsingular matrix S such that $S^{-1}AS=\Lambda$, where Λ is a diagonal matrix containing the eigenvalues of A. Thus,

$$L\sum_{h=1}^{H}(A^{H-h}\otimes A^{h-1})D=L\sum_{h}(S\Lambda^{H-h}S^{-1}\otimes S\Lambda^{h-1}S^{-1})D$$
$$=L(S\otimes S)\sum_{h}(\Lambda^{H-h}\otimes\Lambda^{h-1})(S^{-1}\otimes S^{-1})D$$
$$=L(S\otimes S)DL\sum_{h}(\Lambda^{H-h}\otimes\Lambda^{h-1})DL(S^{-1}\otimes S^{-1})D,$$

by Lemmas 4.4 (i) and 4.5 (v). Since $L(S^{-1}\otimes S^{-1})D=(L(S\otimes S)D)^{-1}$, and using Lemma 4.4 (ii), we have

$$\left|L\sum_{h=1}^{H}(A^{H-h}\otimes A^{h-1})D\right|=\left|L\sum_{h}(\Lambda^{H-h}\otimes\Lambda^{h-1})D\right|.$$

Lemma 4.1 tells us that $L(\Lambda^{H-h}\otimes\Lambda^{h-1})D$ is a diagonal matrix with elements $\lambda_i^{h-1}\lambda_j^{H-h}$, $i\geq j$. Hence,

$$\left|L\sum_{h=1}^{H}(A^{H-h}\otimes A^{h-1})D\right|=\prod_{i\geq j}\left(\sum_{h=1}^{H}\lambda_i^{h-1}\lambda_j^{H-h}\right)=H^n|A|^{H-1}\prod_{i>j}\mu_{ij},$$

with

$$\mu_{ij}=\sum_{h=1}^{H}\lambda_i^{h-1}\lambda_j^{H-h}=\frac{(\lambda_i^H-\lambda_j^H)}{(\lambda_i-\lambda_j)}.$$

If A has multiple eigenvalues, $\lambda_i=\lambda_j$ say, we again consider this as a limiting case of the situation where λ_j approaches λ_i. Taking the limit as $\lambda_j\to\lambda_i$ we find $\mu_{ij}=H\lambda_i^{H-1}$. □

Proof of Lemma 4.6. We shall only consider the determinant of the sum of $L(A\otimes A)D$ and $L(B\otimes B)D$. The determinant of their difference is proved in the same way. By Lemma 4.4 (i),

$$L(A\otimes A+B\otimes B)D=\big(I+L(BA^{-1}\otimes BA^{-1})D\big)L(A\otimes A)D.$$

If BA^{-1} has eigenvalues λ_i, $i=1\cdots n$, $L(BA^{-1}\otimes BA^{-1})D$ has eigenvalues $\lambda_i\lambda_j$, $i\geq j$, by Lemma 4.4 (iii), so that, using Lemma 4.4 (iv),

$$|L(A\otimes A+B\otimes B)D|=\prod_{i\geq j}(1+\lambda_i\lambda_j)|A|^{n+1}.$$

To prove (ii) we first assume that A is nonsingular. Then,

$$|L(A\otimes A+B\otimes B)D|=|A|^{n+1}\prod_{i\geq j}\left(1+\frac{b_{ii}b_{jj}}{a_{ii}a_{jj}}\right)=\prod_{i\geq j}(a_{ii}a_{jj}+b_{ii}b_{jj}),$$

since A and B are now lower triangular. If A is singular, we obtain (ii) starting with $A+\delta I$, where δ is small and $A+\delta I$ is nonsingular.

Consider now case (iii) where $AB=BA$. This result can be proved applying the same method as in the proof of Lemma 4.4 (vii). □

Proof of Lemma 4.7. We shall only show (iii) and (iv), as (i) and (ii) can be proved similarly. Since A is symmetric and nonsingular by assumption, we have from the implicit definition of D and Lemma 4.4, $DL\operatorname{vec}A=\operatorname{vec}A$, $DL(A\otimes A)D=(A\otimes A)D$, $|L(A\otimes A)D|=|A|^{n+1}$, and $(L(A\otimes A)D)^{-1}=L(A^{-1}\otimes A^{-1})D$. Thus,

$$\begin{aligned}L(A\otimes A+\alpha\operatorname{vec}A(\operatorname{vec}A)')D&=\left[I+\alpha L\operatorname{vec}A(\operatorname{vec}A)'DL(A^{-1}\otimes A^{-1})D\right]L(A\otimes A)D\\&=\left[I+\alpha L\operatorname{vec}A(\operatorname{vec}A)'(A^{-1}\otimes A^{-1})D\right]L(A\otimes A)D\\&=\left[I+\alpha(L\operatorname{vec}A)(D'\operatorname{vec}A^{-1})'\right]L(A\otimes A)D.\end{aligned}$$

Since for any two vectors x and y of the same order,

$$|I+xy'|=1+y'x \quad\text{and}\quad (I+xy')^{-1}=I-\frac{xy'}{1+y'x},$$

we find

$$\begin{aligned}|I+\alpha(L\operatorname{vec}A)(D'\operatorname{vec}A^{-1})'|&=1+\alpha(\operatorname{vec}A^{-1})'DL\operatorname{vec}A\\&=1+\alpha(\operatorname{vec}A^{-1})'\operatorname{vec}A=1+\alpha\operatorname{tr}A^{-1}A=1+\alpha n,\end{aligned}$$

and

$$\left[I+\alpha(L\operatorname{vec}A)(D'\operatorname{vec}A^{-1})'\right]^{-1}=I-\frac{\alpha}{1+\alpha n}L\operatorname{vec}A(D'\operatorname{vec}A^{-1})'.$$

Hence,

$$|L(A\otimes A+\alpha\operatorname{vec}A(\operatorname{vec}A)')D|=(1+\alpha n)|L(A\otimes A)D|=(1+\alpha n)|A|^{n+1},$$

and

$$\begin{aligned}&\left[L(A\otimes A+\alpha\operatorname{vec}A(\operatorname{vec}A)')D\right]^{-1}\\&\qquad=L(A^{-1}\otimes A^{-1})D\left[I-\frac{\alpha}{1+\alpha n}L\operatorname{vec}A(\operatorname{vec}A^{-1})'D\right]\\&\qquad=L\left[A^{-1}\otimes A^{-1}-\frac{\alpha}{1+\alpha n}(A^{-1}\otimes A^{-1})DL\operatorname{vec}A(\operatorname{vec}A^{-1})'\right]D\\&\qquad=L\left[A^{-1}\otimes A^{-1}-\frac{\alpha}{1+\alpha n}(\operatorname{vec}A^{-1})(\operatorname{vec}A^{-1})'\right]D.\ \square\end{aligned}$$

REFERENCES

[1] T. W. ANDERSON, *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.

[2] P. BALESTRA, *La dérivation matricielle*, Collection de l'IME, no. 12, Sirey, Paris, 1976.

[3] M. BROWNE, *Generalized least squares estimation in the analysis of covariance structures*, South African Statist. J., 8 (1974), pp. 1–24.

[4] J. CONLISK, *The equilibrium covariance matrix of dynamic econometric models*, J. Amer. Statist. Assoc., 64 (1969), pp. 277–279.

[5] W. L. DEEMER AND I. OLKIN, *The Jacobians of certain matrix transformations useful in multivariate analysis*, Biometrika, 38 (1951), pp. 345–367.

[6] H. V. HENDERSON AND S. R. SEARLE, *Vec and vech operators for matrices, with some uses in Jacobians and multivariate statistics*, Canad. J. Statist., 7 (1979) pp. 65–81.

[7] H. JACK, *Jacobians of transformations involving orthogonal matrices*, Proc. Roy. Soc. Edinburgh Sect. A, 67 (1966), pp. 81–103.

[8] J. R. MAGNUS AND H. NEUDECKER, *The commutation matrix: some properties and applications*, Ann. Statist., 7 (1979), pp. 381–394.

[9] D. G. NEL, *On the symmetric multivariate normal distribution and the asymptotic expansion of a Wishart matrix*, South African Statist. J., 12 (1978), pp. 145–159.

[10] H. NEUDECKER, *Some theorems on matrix differentiation with special reference to Kronecker matrix products*, J. Amer. Statist. Assoc., 64 (1969), pp. 953–963.

[11] I. OLKIN, *Note on 'The Jacobians of certain matrix transformations useful in multivariate analysis'*, Biometrika, 40 (1953), pp. 43–46.

[12] I. OLKIN AND A. R. SAMPSON, *Jacobians of matrix transformations and induced functional equations*, Linear Algebra and Appl., 5 (1972), pp. 257–276.

[13] R. PENROSE, *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc., 51 (1955), pp. 406–413.

[14] C. R. RAO, *Linear Statistical Inference and its Applications*, 2nd ed., John Wiley, New York, 1973.

[15] J. F. RICHARD, *A note on the information matrix of the multivariate normal distribution*, J. Econometrics, 3 (1975), pp. 57–60.

[16] D. S. TRACY AND P. S. DWYER, *Multivariate maxima and minima with matrix derivatives*, J. Amer. Statist. Assoc., 64 (1969), pp. 1576–1594.

[17] D. S. TRACY AND R. P. SINGH, *Some modifications of matrix differentiation for evaluating Jacobians of symmetric matrix transformations*, Symmetric Functions in Statistics, D. S. Tracy, ed., University of Windsor, Windsor, Ontario, 1972.

[18] W. J. VETTER, *Vector structures and solutions of linear matrix equations*, Linear Algebra and Appl., 10 (1975), pp. 181–188.

[19] A. ZELLNER, *An Introduction to Bayesian Inference in Econometrics*, John Wiley, New York, 1971.